

# 作文評価と構文的特徴について —留学生のプレイスメントテストを例に—

長谷川守寿

## 1. 目的

本研究の目的は、プレイスメントテストの一環として行われる作文テストと、その評価の関係を作文の構文的特徴から考察し、作文テスト評価システム構築に活かしていくことである。

プレイスメントテストに作文テストを含めた場合、大量の枚数の評価を短時間に、しかも限られた人数で行うことが求められる。これは評価者への負担が非常に大きく、評価する際の客観的な指標が望まれる。そこで本研究では、作文評価の際には作文を構成する文の構文情報が1つの要因となるのではないかと考え、文の数、従属節の数、格成分の数と評価の関係を考察し、自動作文評価システム作成への一歩としたい。

## 2. 先行研究

まず、作文評価基準案として、森田(1981)・菊池(1987)がある。森田(1981)は、総合的評価としての基準を「初級」と「中級・上級」のレベルに分けて提示している。菊池(1987)は、外国人が日本語の作文を書くための能力は「日本語能力」と「文章能力」に大きく分けられるとし、それらを考慮した方法を提案している。また、作文評価の観点に関する研究としては、田中・坪根・初鹿野(1998)、田中・初鹿野・坪根(1998)がある。田中・坪根・初鹿野(1998)では、日本語教師と一般の日本人が初級から上級までの6レベルの作文を評価した時の観点(22の評価項目)を因子分析し、作文評価の基本構造を示している。さらに、田中・初鹿野・坪根(1998)では、実際の作文評価において作文評価構造のどの要素が重視されるか検討し、「正確さ」などを重視するグループと高度な文型などを重視するグループがあり、ともに「内容」「主旨の明確さ」「構成」が良い作文の決定要因と考えている可能性があることを明らかにした。

これらの問題点として、まず森田(1981)では初めに作文のレベルを固定してから評価を行うので、プレイスメントテストでは使用できない。また菊池(1987)や田中ほか(1998)で挙げている評価基準は、少数の作文評価には適当であるが、大量の作文を評価する際にこれらの観点をを用いるのは困難であり、主観が含まれる余地も多いと思われる。

これらの研究は、教師が学習者の作文能力を測るための評価の観点を明らかにしたものである。一方、教師が学習者の作文を評価するのと同じように自動化システムが評価を行うには、どのような観点が必要かを明らかにする研究があり、高ほか(2001)がその例としてあげられる。高ほか(2001)では、n-gram を利用し、文字列情報と評価の関係を考察している。その結果、日本語学習者の作文(感想文)と対象となる小説の類似度から作文の書き方の特徴が抽出でき、それを日本語らしさの評価指標として使用できること、また同じテーマで書かれた日本人学生の作文との類似度が作文を評価する指標として期待できることを示している。高ほか(2001)で考察している文字列情報の類似度を指標とする手法は、同一レベルの学生が書いた要約文や粗筋文の評価には有効と思われる。しかし、類似度が高くても作文の評価が高くないケースや、類似度は低くても高い評価を得た例もあげられており、文字列のようにきわめて表層的な情報だけ用いることの限界も示している。また学生が教師側の示したテーマで作文を書く、いわゆる自由作文の評価の際には、比較対象となる小説もなく、同じテーマで書かれた日本人の感想文も必要になるため、実際に稼働させるのは難しいと思われる。以上のように、この方法をプレイスメントテストの一環として行う自由作文の評価に使用するには、不適切であると考えられる。

長谷川(2004)では、プレイスメントテスト内の作文と評価の関係について考察する中で、延べ語数、異なり語数は、評価と高い相関関係があることを示した。また品詞数と作文評価では、助詞・動詞・名詞がより高い相関を持つことが分かったが、さらに評価基準として関連性の高い項目を求める必要がある。

### 3. データ

#### 3. 1. 対象データ

筑波大学留学生センターで実施されたプレイスメントテストの中の、作文テストを対象とする。プレイスメントテストは、SPOT<sup>註1</sup>・マークシートテスト・作文テストからなっており、留学生センターで開講している授業の履修を希望する留学生は全てこのテストを受けなければならない。1999年2学期に実施された作文テストは、筑波大学の学生の起床時間と就寝時間に関する調査結果(参照:資料左:元のデータは筑波大学新聞第181号1997年5月19日発行による)を見て作文を書いたもの67人分(以後「99年作文」と呼ぶ)で、2002年1学期に実施された作文テストは、4コマ漫画(参照:資料右:フジ三太郎)を見て説明する課題112人分(以後「02年作文」と呼ぶ)である。対象者の国籍、日本語歴、レベルは様々である。全数調査を予定し

たが、一部欠けているデータもあったので、対象となったのは、SPOT・マークシートテスト・作文テスト全てを受験した学生のもので、99年作文は66人分、02年作文は106人分である<sup>注2</sup>。

作文課題は、A4用紙に印刷され、表面には「いろいろな文型を使うこと Use a variety of grammatical patterns. 知っている漢字は、たくさん書くこと Use Kanji as appropriate.」という毎年共通の指示があり、裏面には資料にあるような各年個別の指示がある。作文の制限時間は15分である。

### 3. 2. データのチェック

本研究で使用する作文テストの原稿は、全てテキスト化されたものである。本研究では、形態素解析という機械処理を施した場合、同様の結果が得られるように、データの一貫性に関するチェックを行った。これは、長谷川(2004)と同様に、基準が統一されるように筆者が行った。

### 3. 3. 構文解析ツール

本研究で直接集計の対象となるデータは、前述の方法でテキスト化され、チェックを受けた作文を、構文解析ツール CaboCha(工藤ほか(2002))で処理したものである。CaboChaは、例えば「田中さんは初めて朝はやくおきて公園でジョギングをした。」(99年作文より)という文に対して、形態素解析システムである茶釜(松本ほか(1999))と同様、(1)のように形態素解析を行うことも可能であるが、(2)のように係り受けを示した木構造や、(3)のような文節区切りをすることも可能である。

(1)	田中	タナカ	田中	名詞・固有名詞・人名・姓	B-PERSON	
	さん	サン	さん	名詞・接尾・人名		0
	は	ハ	は	助詞・係助詞		0
	初めて	ハジメテ	初めて	副詞・一般		0
	朝	アサ	朝	名詞・副詞可能		0
	はやく	ハヤク	はやい	形容詞・自立		
				形容詞・アウオ段	連用テ接続	0
	おき	オキ	おきる	動詞・自立一段	連用形	0
	て	テ	て	助詞・接続助詞		0
	公園	コウエン	公園	名詞・一般		0
	で	デ	で	助詞・格助詞・一般		0

ジョギング	ジョギング	ジョギング	名詞一般	0		
を	ヲ	を	助詞-格助詞一般	0		
し	シ	する	動詞-自立	サ変・スル	連用形	0
た	タ	た	助動詞	特殊・タ	基本形	0
。	。	。	記号-句点			0

(一部見やすいように著者による修正済み)

(2) <PERSON>田中</PERSON>さんは-----D

初めて----D |  
朝-D | |  
はやく-D | |  
おきて----D  
公園で---D  
ジョギングを-D  
した。

(3) \* 0 7D 1/2 3.39770032

田中	タナカ	田中	名詞-固有名詞-人名-姓	B-PERSON
さん	サン	さん	名詞-接尾-人名	0
は	ハ	は	助詞-係助詞	0

\* 1 4D 0/0 0.94183603

初めて	ハジメテ	初めて	副詞一般	0
-----	------	-----	------	---

\* 2 3D 0/0 0.23730092

(以下、略)

しかし、表記や文法に関する誤りが多く、また通常漢字で表記される文字がひらがなで記述されることが多い留学生の作文を解析した場合、結果に誤りを多く含むことになってしまう。それでも、作文評価の自動化を想定した時、形態素解析の結果を人手で修正する処理を加えることは自動化という考え方に反するし、長谷川(2004)で明らかにしたように、形態素解析の結果への修正の有無には有意な差がないことから、本研究でも解析結果をもとに構文情報を抽出する際には、修正は行わずに進めることとする。

### 3. 4. 作文の評価について

本調査の前提となる作文の評価について説明する。作文の評価は、日本語教育歴が

長くさらに作文教育担当の経験がある教師によって行われ、0から6までの7段階でつけられている。

99年作文の評価者は1名であるが、02年作文の評価者は2名である。そこで、まず02年作文に対して2名が下した評価の相関に関して、評定者間の信頼性を調べたところ、相関係数は.832であった(スピアマンの順位相関係数、危険率0.1%で有意)。説明率は69.0%で、2名の評価者間には強い相関があるといえる。

なお、評価者の判定が一致したデータのみ使用、または判定を修正して使用することも考えられるが、一致しているデータは106データ中36データと少なく、評価1が0、評価2が16、評価3が4と評価ごとのばらつきも多い。そのためこの方針はとらず、02年作文に関しては2名の評価者を使用して、99年作文と02年作文というテーマによる違いと、2名の評価者(以後評価者A、Bと呼ぶ)による個人差も考察する。

#### 4. 方法

本研究では、自動作文評価システム・作文評価補助システムを目標とした研究の一環として、機械的な作文評価の方法を考察する。作文評価の手掛かりとしては、作文の表層情報と深層情報が考えられる。表層情報には単純にどのような文字列からなるかという文字列情報と、どのような形態素からなるかという形態素情報が考えられ、深層情報には、構文・意味情報が含まれると考える。長谷川(2004)では、表層情報に着目し、形態素情報と評価の関係について考察した。本研究では、作文の評価は作文の構文的特徴と関係があるのか調べるために、文の深層構造、その中でも構文情報(含まれる文の数、格成分の数、従属節の数)に注目し、作文における構文情報と評価の関係を考察する。

本研究では、作文の構文的特徴という観点から考察を行う。これは文がたくさん書いていけば評価が高くなっているのか、従属節が加わることによって単文よりも構造が複雑な複文が書いていけば、評価が高くなっているのか、格成分が多く、色々な要素を含んでいた方が評価が高いのか、それぞれを考察するためである(なお、省略等の問題は意味解析では重要な要因であるが、今回は構文情報であるため触れない)。

また、システム構築の方針として、使用するフリーのソフトウェアの解析結果に多少誤りが含まれていても、それらを使用して構築することとしたい。これは、全てをゼロから作り出すことが不可能であるということもあるが、長谷川(2004)で検討したように、形態素解析結果に修正を加えることに、大きな効果が見られなかったため、誤りを修正することのメリットは少ないと考えるからである。

#### 4. 1. 文のカウント方法

文のカウントの方法は計量的研究では常に問題となる点であるが、本研究では作文に含まれる文のカウント方法は、単純に句点「。」を数えるものとし、引用文中に句点が含まれていても、1文とカウントする(なおこれ以降、作文に含まれる文の数を「文数」と呼ぶこととする)。ただし、日本語レベルの低い学習者の作文には、(4)のように文が途中までしか書けていないものもあった。これは15分という時間制限があったためと、述語が思いつかない等の理由で、文を最後まで完成することができなかったためと思われるが、本研究では途中までしか書けていないものは、1文とは認めずにカウントしないものとする。なお、こういった要素をカウントするスクリプトは、Perlを使用して記述した。

(4) 4じから6じまでおきる学生は (99年作文より)

また、学習者の作文の中には句点が全く使われておらず、次の文との間にスペースだけがおかれているようにみえるものや、句点の使い方が分かっておらず読点「、」と混同して使われているものがある。対象データには、99年作文で3例、02年作文で4例あったが、(5)のようなものは文とは認めず、句点を持つもののみ文と認めた。

(5) (略) いっしょにはやいがっこへいきましょ、ちょっとまってください  
(02作文より)

但し、句点をカウントすることによって、以下の問題が考えられる。まず、引用文の中の問題である。99年作文に1例、02年作文の8例あるのだが、(6)(7)のように引用文中に句点が使われているため、1文とカウントしてしまうという問題が起きる。これはCaboChaでは厳密な文末区切りができないために起こる問題であり、3.3で述べたような理由から修正は行わないこととする。

(6) (略) ……の学生に「あなたは普段何時に寝ますか。そして何時に起きますか」という調査を…… (99年作文より)

(7) 「今度はまけるはずがない。僕がもう若いんじゃないか。」と思って  
(02年作文より)

またこの方法でカウントする場合、感嘆符「!」、疑問符「?」を文末に持つ例が排除されることになる。しかし、全体への影響は少ないと判断できる(99年作文に1例、02年作文に5例)ことと、プレイスメントテストの作文には、「???になって、町でねいました」(02年作文より)のように、受験者が言葉が分からない部分に使うこともあり、文末を示す記号としては排除した。

#### 4. 2. 従属節数のカウント方法

従属節数のカウント方法であるが、本研究では、接続助詞 (CaboCha では「から・が・けど・し・ちゃ・て・で・と・ながら・ので・のに・ば・ものの」) の数に、動詞・形容詞の連用テ形・連用形+読点「、」, 断定の助動詞「だ」の連用形「たら」「ても」の数値を加えた。なお、CaboCha では、形容動詞という分類はされておらず、名詞+断定の助動詞 (CaboCha では「名詞・形容動詞語幹」と「助動詞 特殊・ダ」) に分かれて解析される。そこで、「助動詞 特殊・ダ」の部分が現れた段階でカウントした。なお、動詞の連用テ形は、補助動詞 (CaboCha では動詞-非自立と解析される) に繋がる場合には、カウントしない。例えば、「食べている」などの「食べて」の部分は、カウントしないこととなる。

なお、すべてにおいて CaboCha のシステム上の誤りには、手を加えていない。例えば、(8)の下線部のような表現は、CaboCha では断定の助動詞「だ」の連用形と解析できない (格助詞と解析される) ので、従属節として検出するのは不可能である (なお、(9)のような場合は正しく解析できる)。また、接続助詞「と」にもシステム上の誤りがあるが、これも同様の対応をとった。

(8) 1時が45%以上で最高だった。 (99年作文より)

(9) 1時が45%以上で、最高だった。 (著者作例)

この方法では、(10)のように連体修飾節内の従属節 (「昼をすぎて」の部分) も「動詞連用テ形+読点」としてカウントされるが、出現数はごくわずかで、影響は少ないものと考えられる。

(10) 昼をすぎて、13時や14時に遅く起きる人もいました。 (99年作文より)

ちなみに、この方法で(11)の文に含まれる従属節をカウントした場合、「……説明やすいですが」「……はじまるので」「……起きなければならなかったら」「……寝ぼうしていけないので」「……一時ぐらい寝るなら」の合計5となる。

(11) おきる時間のグラフは説明やすいですが筑波大学の学生達は毎朝はじめてじゅうぎょうのために八時ぐらいに起きます。じゅうぎょうは八時四十分はじまるので八時半に起きるべきです。もしみんな七時半から八時半まで起きなければならなかったら自分の体によって七時間とか八時間ぐらいねるのが必要です。それで朝寝ぼうしていけないので一時ぐらい寝るなら寝る時間が足りません。

(99年作文より)

CaboCha は当然のことながら正しい文が入力されてくることを想定して作られているので、日本語として正しくない文が入力された時、問題がおこる。例えば、(12)

の場合、「いきましたので」の部分は日本語として正しくないので、解析結果も誤ったものとなり、結果として接続助詞が含まれているとされないため、カウントしない。

(12) 土・日は友達と一緒にあそびにいきましたので、 (以下略。99年作文より)

#### 4. 3. 格成分数のカウント方法

本研究で使用した CaboCha の機能には、文節区切りが含まれるので、文節区切りされたものの中で「名詞+格助詞」「名詞+係助詞」と連続するものの数をカウントした。なお、名詞には接尾辞に相当するもの (CaboCha では名詞-接尾と解析されるもの) も含まれる。たとえば、「田中さん」の「さん」は「名詞-接尾-人名」と解析される。また、日本語学習者の作文の場合、「名詞+係助詞」は「名詞+格助詞」を兼務している場合が多いと考え、格成分数として数えた。例えば(13)の「田中さんは」はガ格を兼務していると考えられる。このように数えたとき、(13)に含まれる格成分は、「田中さんは」「公園で」「ジョギングを」の3つである。

(13) 田中さんは初めて朝はやくおきて、公園でジョギングをした。(99年作文より)

なお、兼務していないものとしては、(14)のように日本語として不適切なものもあるが、これらもカウントすることとなる。

(14) そして25%の学生は生舌に不明だ。(99年作文より)

また、(15)の下線部のように助詞が付加していないものはカウントしない。

(15) (略)、筑波大学の学生、よるあそぶで、朝、(略) (99年作文より)

## 5. 結果

表1 99年作文と02年作文の構文情報(99年はN=66、02年はN=106)

	99年	02年
文数	7.8	8.9
	3.2	3.6
従属節数	1.9	2.9
	1.6	1.9
格成分数	19.5	20.2
	8.9	7.9

(上段が平均値、下段が標準偏差)

作文データを構成する個々の作文の書き手が異なり、書き手の構成も異なるため、



単純に比較できないが<sup>註3</sup>、データ全体を比較した時、表1のような結果となった<sup>註4</sup>。

この2つのデータを比較する限り、文数・従属節数・格成分数において、99年作文より02年作文の方が多いたことが指摘できる。99年作文の課題は、表に数値が記載されていて情報が多く、また細かい指示があるため、99年作文の方が文をたくさん書け、従属節を使った複雑な構文を使用し、格成分をたくさん持ち色々な要素を含んだ文が書けるのではないかと予想したのだが、結果としては変化のある4コマ漫画の方が3つの観点全てで数値が上回った。今回はグラフを説明する課題と4コマ漫画を説明する課題の比較だけであるが、更に課題と書かれる文の関係を調べる必要がある。

### 5. 1. 文数と作文評価の相関

まず、評価と文数の関係を調べる。一例として、99年作文の評価と文数の散布図を図1に示す。

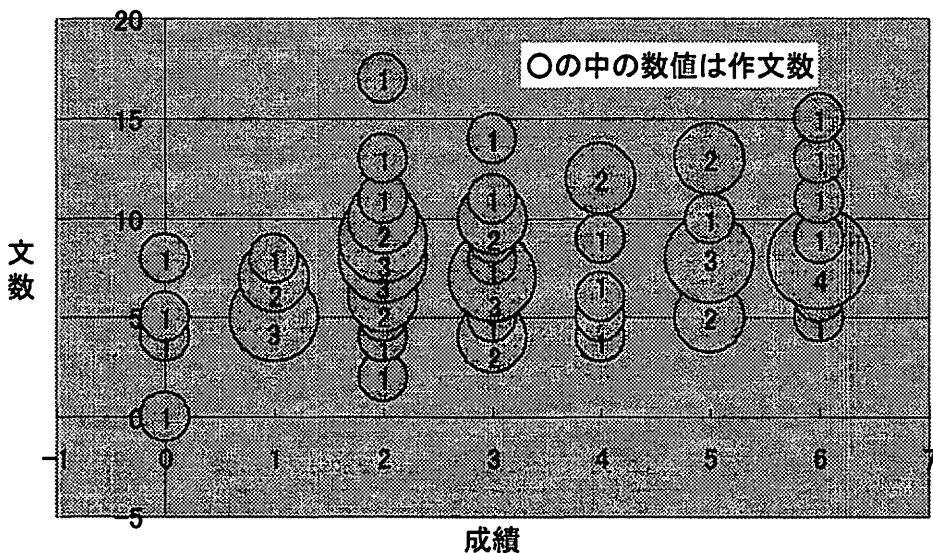


図1. 文数と作文評価の散布図(99年)

評価2で文数が17文というはぐれ値もあるが<sup>註5</sup>、評価が上がるに従って、文数がかかる範囲も上昇しているように見える。そこで、作文の評価と作文の文数の相関関係を、スピアマンの順位相関係数により調べると、99年作文と評価の相関係数は  $r_s=0.280$  (5%で有意)、02年作文と評価の相関係数は、評価者Aの場合  $r_s=0.192$  (5%で有意)、評価者Bの場合  $r_s=0.182$  (有意差なし) と、「ほとんど相関がない」か「やや相関がある」と判断できる程度である。長谷川(2004)で示されたように動詞の数、名詞の数と

評価の相関は高かった（99年作文の修正ありデータでは、動詞・名詞の数と評価には「強い相関がある」）が、それらが述語となっている文数と評価の相関はそれほど高くなく、自動評価の観点としてはあまり有効ではないものと考えられる。

本研究では、動詞の連体修飾用法に対する調査はしていないが、作文中には「寝る時間」「起きる人」などの表現も見られ、連体修飾用法を今後調査対象とすることも必要と考えられる。

## 5. 2. 従属節数と作文の評価の関係

ここでは、作文の評価と、従属節数との関係を見る。例として、99年作文の評価と格成分数との散布図を図2に示す。

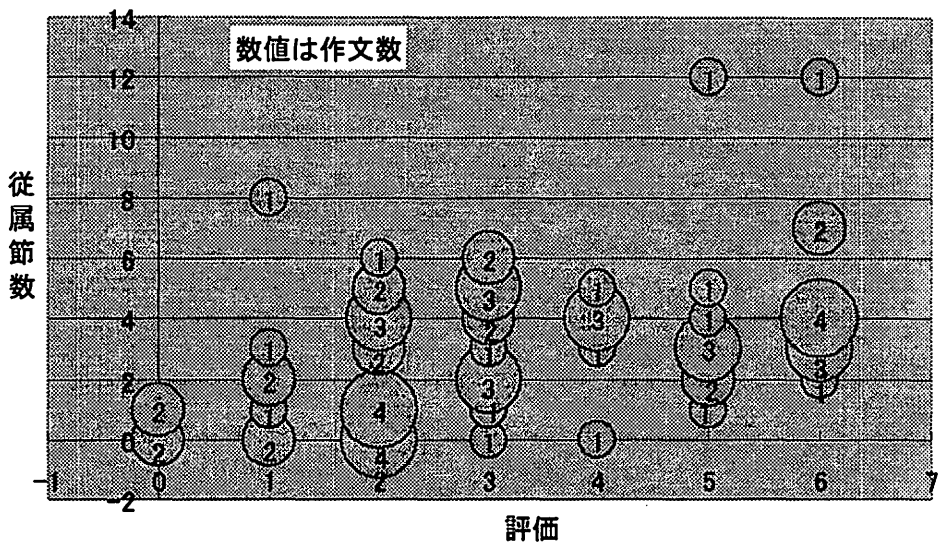


図2 99年作文の評価と従属節数の散布図

評価が上がるに従って、従属節数がかかる範囲も上昇しているので、5. 1と同様に従属節数と作文評価の相関関係を調べると、99年作文と作文評価の相関係数は  $r_s = .36$ （1%で有意）、02年作文と評価では、評価者 A との場合  $r_s = .54$ 、B との場合は  $r_s = .56$ （ともに1%で有意）で、「やや相関がある」「相関がある」という結果になった。これを見ると、動詞、助詞の数と評価は「強い相関がある」関係だったが、それらの組み合わせである従属節と評価の相関はそれほど高くないといえる。また課題による違いも関係していると思われる、4コマ漫画のようなものでは、従属節を用い、並列構造や論理構造を持った複雑な表現の数と評価の相関がやや高いようである。また現在は接続助詞「て」も接続助詞「のに」も同じようにカウントしているが、導入時期の異

なる表現であり、難易度も異なると思われる。今後はどのような接続表現が使われるのか、といった観点からの調査も必要であろう。

### 5. 3. 格成分数と作文の評価の関係

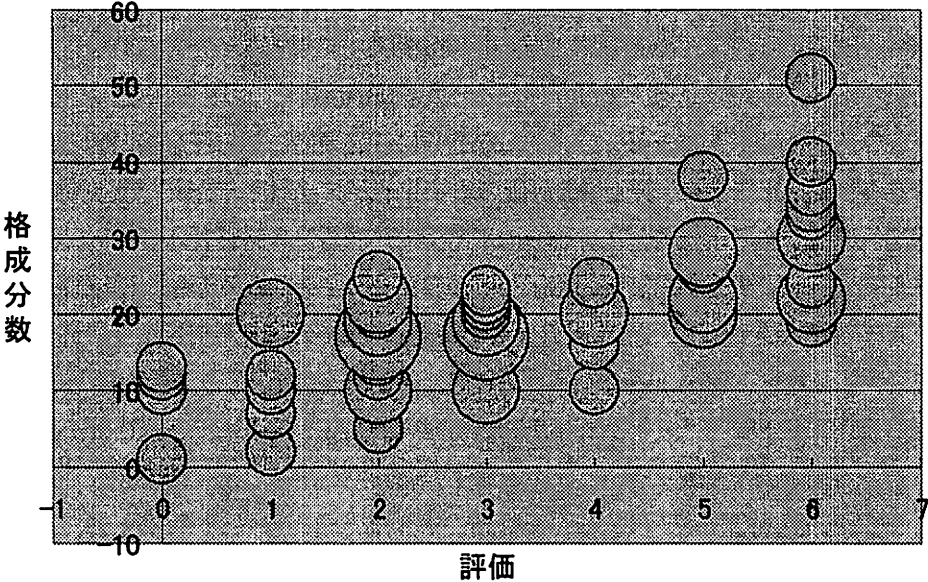


図3. 99年作文の評価と格成分数の散布図

ここでは作文の評価と格成分数の関係を見る。例として99年作文の評価と格成分数の散布図を図3に示す。なお図3に関しては、図が煩雑になるため作文数の数値を省略した。小さい丸が1つの作文、それ以上の大きさの丸は2つ以上の作文を表す。

図3より評価が上がるに従って、格成分数をとる範囲も上昇していることが確認できる。そこで、作文の格成分数と作文の評価の相関関係を調べると、99年作文と評価の相関係数は  $r_s = .71$  (1%で有意) で「強い相関がある」、02年作文の格成分数と評価者Aの評価との相関係数は  $r_s = .46$  (1%で有意) で「相関がある」、02年作文の格成分数と評価者Bの評価との相関係数は  $r_s = .56$  (1%で有意) で「相関がある」である。

全体として、名詞と助詞数と評価の相関は高かったが、それらの組み合わせである格成分数と評価の相関はそれ程高くなかったといえる。これは、名詞の副詞的用法(「朝8時にねる」の「朝」や、「毎日2時間寝る」の「毎日」)が調査の対象から外れたことも影響しているのではないかと考えられる。

## 6. まとめ

長谷川(2004)では、助詞・動詞・名詞が評価と高い相関を持つことが分かっている。しかし、本研究で想定した構文情報と評価の間には、「強い相関がある」ことを示すものはなかった。まず、文数は「ほとんど相関がない」または「やや相関がある」程度であった。従属節数は、02年作文の場合、評価と「相関がある」関係であり、格成分数は、99年作文の場合、評価と「強い相関がある」結果となった。これは、作文課題による違いが考えられるのではないだろうか。99年作文では限られた情報（「ねる」「おきる」「時間」「パーセンテージ」など）に、必須格にしろ任意格にしろ色々な成分をつけていくことに、評価者の評価のポイントがあり、02年作文では並列関係や論理関係をはっきり示すことに評価者のポイントがあると考えられる。このように課題によって評価者の求める評価のポイントが異なり、その異なりが結果に反映したものと考えられるのである。このように考えた時、学習者の能力を測るプレースメントテストの作文には、どのようなタイプの課題を出すのが望ましいのかを明らかにする必要がある。また、これらの評価の観点を組み合わせて、人手による作文の評価に近い数値を導く方法も考えていくことも考えたい。

## 7. 今後の課題

本研究では、自動作文評価システム構築に向け、深層情報、特に作文の構文情報を手掛かりとして、評価との関連を考察してきた。構文情報は、数量的側面のみで、語の難易度などは考慮していない。例えば5.2で見たように接続助詞「て」も接続助詞「のに」も同じ1つのものとして数えている。そこでより高い相関関係を持つ項目を探すためには、日本語能力試験の出題基準などを参考に、作文に使われている語彙の難易度の情報と作文の評価の関係を考察することも重要になってくるかと思われる。

また、今回使用したシステムは日本語として正しいもののみをカウントしている。しかし、採点者は実際には誤りを補正しながら読んでいると考えられる。例えば、02年作文に出てくる「ジョギング」には、「ジグンク・ショキング・ジョーキング・チョキング」など色々な誤りがあるが、これらの語は辞書に登録されていないため未知語とされ、格成分数には反映されない。誤りを修正しないという立場で行ってきたが、語の類似度から未知語の推定を行い、評価者が行っているのと同じような過程を踏む方法も考慮していきたい。

謝辞：本研究で使用したデータは、筑波大学留学生センター西村よしみ教授からお借りしたものである。感謝申し上げます。

## 注

- (1) SPOTは、フォード丹羽順子ほかが開発した（フォード丹羽ほか(1995)）。
- (2) フェースシートによると、学生の国籍は、99年作文では、66人中、中国・26人、韓国・17人、アメリカ・10人、台湾・4人、マレーシア・3人、スロベニア・2人、オーストラリア・スリランカ・ブラジル・メキシコ・各1人で、02年作文では、106人中、中国・42人、韓国・35人、タイ・4人、ハンガリー・ペルー・各3人、アメリカ・マレーシア・ロシア・各2人、アゼルバイジャン・イギリス・オーストラリア・カナダ・スリランカ・セネガル・チリ・パキスタン・フィリピン・ブラジル・フランス・メキシコ・台湾・各1人である。
- (3) 作文の評価者によって、作文は表Aのように評価となった。99年作文には、評価0の作文が、4つあったことを示す。

表A. 作文データにおける評価の構成

評価	0	1	2	3	4	5	6	合計
99年作文	4 6%	7 11%	16 24%	13 20%	6 9%	9 14%	11 17%	66 100%
02年作文A	0 0%	2 2%	32 30%	24 23%	18 17%	20 19%	10 9%	106 100%
02年作文B	2 2%	3 3%	19 18%	12 11%	20 19%	21 20%	29 27%	106 100%

- (4) 分析には柳井久江『4Steps エクセル統計』(1998年、オーエムエス出版)内のStatcelを使った。
- (5) (a)は、この学生が書いた作文の一部である。テーマと直接関係のない文を書いているため、このような結果になったものと思われる。
  - (a) わたしは筑波大学の学生です。あなたたちも筑波大学の学生です。わたしたちの学校は茨城県います。東京と近くにいます。(以下省略)
- (6) 相関関係の強さの目安には色々異なる種類のものがあるが、柳井(1998,p164)に従う(以下の表bに示す)。

表 b. 相関関係の強さのめやす

相関係数	相関関係
0.0 ~ ±0.2:	ほとんど相関がない
±0.2 ~ ±0.4:	やや相関がある
±0.4 ~ ±0.7:	相関がある
±0.7 ~ ±0.9:	強い相関がある
±0.9 ~ ±1.0:	極めて強い相関がある

### 参考文献

- 菊池康人(1987)「作文の評価方法についての一私案」『日本語教育』第 63 号、87-104
- 工藤拓・松本裕治 (2002)「チャンキングの段階適用による日本語係り受け解析」『情報処理学会論文誌』、Vol.43、No.6、1834-1842
- 高建斌・小高知宏・小倉久和(2001)「外国人日本語学習者による日本語作文の n-gram モデルを用いた特徴抽出と作文評価」『福井大学工学部研究報告』、Vol.49、No.2、217-225
- 田中真理・初鹿野阿れ・坪根由香里 (1998)「第二言語としての日本語における作文評価—『いい』作文の決定要因」『日本語教育』第 99 号、60-71
- 田中真理・坪根由香里・初鹿野阿れ(1998)「第二言語としての日本語における作文評価基準—日本語教師と一般日本人の比較—」『日本語教育』第 96 号、1-12
- 長谷川守寿(2004)「プレイスメントテストにおける作文の評価と形態的特徴の関係について」、筑波大学留学生センター『日本語教育論集』第 19 号、1-20
- フォード丹羽順子・小林典子・山元啓史(1995)「日本語能力簡易試験 (SPOT)」は何を測定しているのか—音声テープ要因の解析—」『日本語教育』第 86 号、93-102
- 松本裕治・北内啓・山下達雄・平野善隆・松田寛・浅原正幸(1999)「日本語形態素解析システム『茶筌』version2.0 使用説明書第 2 版」NAIST Technical Report NAIST-IS-TR99012

(はせがわ もりひさ・首都大学東京准教授)

資料. 作文課題

99年作文

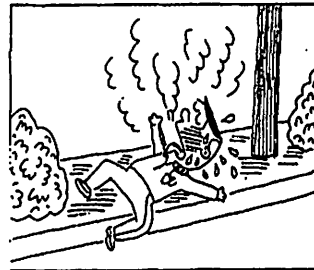
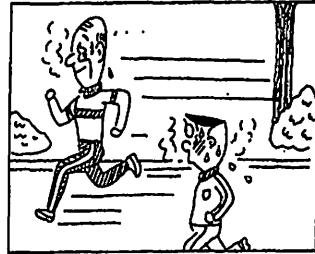
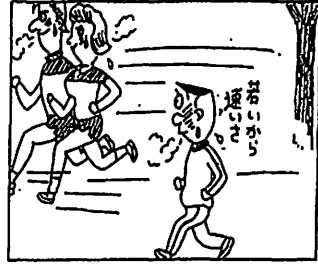
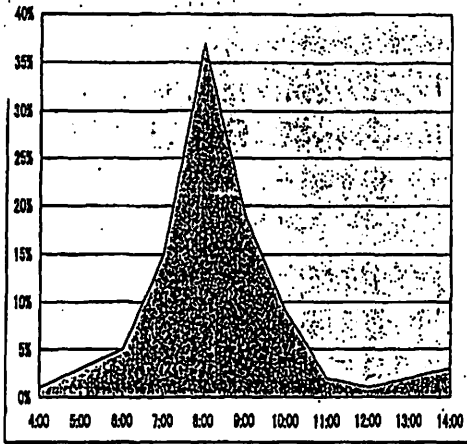
02年作文

下のグラフは、筑波大学の学生150人に「おきる時間」と「ねる時間」について聞いたものである。よく見てグラフを説明しなさい。

題 (topic) 絵を見て説明しなさい。Explain the comic strip

若い わかい  
速い はやい

「おきる時間」



「ねる時間」

